

### STAT100 Problem Set 9: Hypothesis Testing Involving Two Samples

You need to submit a Word document or PDF for this assignment. Make sure you do the following:

1. Upload only one file/document in ELMS for Problem Set 9
2. Include your name in the document in the upper left-hand corner. Under your name, write STAT 100 and your section number. Write Problem Set 9 centered on the page.
3. Number and letter your answers to the questions accordingly.
4. Carefully read all problems and follow all instructions.
5. Upload the assignment in ELMS before the deadline of Sunday 5/8 at 11:59 PM otherwise it is considered late. Make sure you save your document on your computer or email it to yourself so that you keep an electronic copy.

**Students should refer to the Tutorial for Problem Set 9 as they are working on this problem set.** All Tutorials for Problem Sets can be found in the STAT100 ELMS course under Modules.

Problem Set 9 has two questions worth 25 points. **There is no data file for this assignment; all the necessary information and data are provided.**

**You should NOT use RStudio for Problem Set 9.**

1. (11 points) A study examined language acquisition in learning-impaired children. The researcher designed two computer modules that taught children the same content but used different instructional approaches. The study refers to the two computer modules as “Module 1” and “Module 2.” Children were assigned at random to either Module 1 or Module 2. At the end of each module, students took a quiz asking the children to identify 25 words. The number of correct identifications by the children in each Module group were recorded. Data are summarized in the table below (for example, the 175 students who were assigned to Module 1 had a sample mean of 18.07 correct identifications with a sample standard deviation of 4.53).

		Number of correct identifications	
	n	$\bar{y}$	s
Module 1	175	18.07	4.53
Module 2	175	16.92	4.98

Use the provided information to perform a two-sample t-test to determine whether there is a difference between the mean number of correct identifications for students using Module 1 and the mean number of correct identifications for students using Module 2. There is *no* information suggesting that either module should result in a higher (or lower) number of correct identifications.

- State the null and alternate hypotheses. *Be sure to use correct notation and symbols.*
- Is this a two-sided test or a one-sided test? *Explain why.*
- You can assume that the conditions are met for this hypothesis test. Calculate the test statistic for this hypothesis test using the appropriate formula. **YOU MUST SHOW ALL WORK IN YOUR DOCUMENT TO RECEIVE FULL CREDIT.** *Round your final answer to three decimal places.*
- Determine the p-value for this hypothesis test using the STAT100 Z Table. *Use correct notation and state the p-value as a decimal rounded to four decimal places.*
- Draw a conclusion for this hypothesis test and give the reason for your conclusion. *Use a level of significance of  $\alpha = .05$ .*

For question #2, you will use data from a different study.

2. (14 points) Test preparation companies claim that their coaches can raise SAT scores of high school students. A random sample of 310 students who took the SAT twice included 154 students who were not coached before their first SAT attempt but were coached before their second attempt. The sample also included 156 students who were uncoached for either attempt. Students who retake the SAT without coaching typically raise their scores, so we can test the claim that coaching can improve a student's SAT score. The table below summarizes gains in SAT Verbal scores between the first and second attempt for each group of students (for example, the sample mean gain in SAT Verbal scores for the 154 students who were coached after their first attempt was 26.51 points, with a sample standard deviation of 9.67).

	n	Gain in SAT Verbal scores (in points)	
		$\bar{y}$	S
Coached	154	26.51	9.67
Uncoached	156	24.98	9.68

Use the provided information to perform a two-sample t-test to determine whether the mean gain in SAT Verbal scores is *higher* for students who were coached after their first attempt.

- State the null and alternate hypotheses. Use  $\mu_1$  to represent the population mean gain in SAT Verbal scores for the "Coached" group and use  $\mu_2$  to represent the population mean gain in SAT Verbal scores for the "Uncoached" group. *Be sure to use correct notation and symbols.*
- Is this this this a two-sided test or a one-sided test? *Explain why.*
- You can assume that the conditions are met for this hypothesis test. Calculate the test statistic for this hypothesis test using the appropriate formula. **YOU MUST SHOW ALL WORK IN YOUR DOCUMENT TO RECEIVE FULL CREDIT.** *Round your final answer to three decimal places.*
- Determine the p-value for this hypothesis test using the STAT100 Z Table. *Use correct notation and state the p-value as a decimal rounded to four decimal places.*
- Draw a conclusion for this hypothesis test and give the reason for your conclusion. *Use a level of significance of  $\alpha = .05$ .*
- Draw a conclusion for this hypothesis test and give the reason for your conclusion but use a level of significance of  $\alpha = .10$ .
- Did you have the same conclusion for this hypothesis test under the two different levels of significance (for  $\alpha = .05$  and  $\alpha = .10$ )? If not, **provide the value of a test statistic** (not a p-value) that would result in the same conclusion for this hypothesis test under the two different levels of significance.